

DYNAMICAL FAMILIES OF QUADRATIC POLYNOMIALS IN FINITE FIELDS OF CHARACTERISTIC TWO

ERIC BACH AND ANDREW BRIDY

ABSTRACT. Let $k = \mathbb{F}_{2^n}$. Let $f(x) = \alpha x^2 + \beta x + \gamma$ be a map from k to itself with $\alpha, \beta, \gamma \in k$, $\alpha \neq 0$. We show that f is conjugate by a linear polynomial in $k[x]$ to a map in one of two one-parameter families. We then show that the number of equivalence classes of these f under conjugation by any permutation of k is $2^{O(\frac{n}{\log n})}$. In doing so we prove a more general result about affine-linear maps on finite dimensional vector spaces over finite fields.

1. INTRODUCTION AND CHANGE OF VARIABLES

Let k be the field \mathbb{F}_q . We study the dynamical behavior of quadratic polynomial maps from k to k . If q is odd, any quadratic polynomial can be put in the form $x^2 + c$ by completing the square (via conjugation by a linear polynomial in $k[x]$) [16]. If q is even, one cannot complete the square. In this case the corresponding result is the following:

Theorem 1. *Let $k = \mathbb{F}_{2^n}$, and let $f : k \rightarrow k$ be defined by*

$$f(x) = \alpha x^2 + \beta x + \gamma$$

where $f(x) \in k[x]$, $\alpha \neq 0$. Fix any $\zeta \in k$ of absolute trace 1. Then there exists $\psi(x) = ax + b \in k[x]$ such that $\psi^{-1} \circ f \circ \psi(x)$ equals one of the following:

(1) $x^2 + \beta x$

(2) $x^2 + \beta x + (\beta^2 + 1)\zeta$

Moreover, no two distinct maps among those in families (1) and (2) are conjugate by a linear polynomial in $k[x]$.

Proof. First we scale α to be 1. Let $\phi(x) = \alpha^{-1}x$. Then

$$(\phi^{-1}f\phi)(x) = x^2 + \beta x + \alpha\gamma.$$

Date: August 17, 2012.

2010 Mathematics Subject Classification. Primary 37P05, 37P25; Secondary 11T55.

Key words and phrases. Quadratic Dynamics, Finite Fields.

Now let $\psi(x) = \phi \circ (x + c)$ for some c to be determined. Then

$$(\psi^{-1}f\psi)(x) = x^2 + \beta x + c^2 + (\beta + 1)c + \alpha\gamma.$$

If we can find a b such that $c^2 + (\beta + 1)c + \alpha\gamma = 0$ then our map is of the form (1). First assume $\beta \neq 1$ and let $y = \frac{c}{\beta+1}$. We want

$$(y^2 + y)(\beta^2 + 1) = \alpha\gamma$$

which we write as

$$\sigma(y) - y = \frac{\alpha\gamma}{\beta^2 + 1}$$

where $\sigma : x \mapsto x^2$ is the Frobenius automorphism. By the additive version of Hilbert's Theorem 90 [8], there exists a solution if and only if $\text{tr}(\frac{\alpha\gamma}{\beta^2+1}) = 0$, where tr denotes the absolute trace.

Suppose instead that $\text{tr}(\frac{\alpha\gamma}{\beta^2+1}) = 1$. Pick any $\zeta \in k$ of trace 1. Then $\text{tr}(\frac{\alpha\gamma}{\beta^2+1} + \zeta) = 0$, so there exists y such that $y^2 + y = \frac{\alpha\gamma}{\beta^2+1} + \zeta$. Therefore $c^2 + (\beta + 1)c + \alpha\gamma = (\beta^2 + 1)\zeta$, and our map is of the form (2).

If $\beta = 1$, then $(\psi^{-1}f\psi)(x) = x^2 + x + c^2 + \alpha\gamma$, and we choose c such that $c^2 = \alpha\gamma$ and $(\psi^{-1}f\psi)(x) = x^2 + x$. The map $x \rightarrow x^2 + x$ lies in both families (1) and (2).

Uniqueness is immediate. The above trace condition shows that with the exception of the case when $\beta = 1$, no map in family (1) is conjugate to a map in family (2) by any $\psi(x) = ax + b$. Within each family, each map is uniquely parametrized by β , and a conjugation by $ax + b$ does not change β . \square

Theorem 1 shows that, up to conjugation by a linear polynomial in $k[x]$, there exist two one-parameter families of quadratic polynomial maps from \mathbb{F}_{2^n} to itself. This can be thought of as a kind of moduli space for quadratic maps, given explicitly by the intersecting curves $(1, \beta, 0)$ and $(1, \beta, (\beta^2 + 1)\zeta)$ in \mathbb{A}_k^3 .

Remark 2. If we allow conjugation by linear polynomials in $\bar{k}[x]$, every quadratic map lands in family (1). In other words, each type (2) map has a nontrivial \bar{k}/k twist of type (1) as defined in [16]. However, for our purposes we will not extend the field k . One reason for staying in $k[x]$ is the following. If a quadratic map is used in a practical situation, such as part of a pseudo-random generator, we may not be free to change the field of definition. For example, in choosing k for use in a mobile device, there are often severe performance constraints, leading us to want k to be as small as possible.

It may occur that two quadratic maps from k to k are conjugate by a map other than a linear polynomial. We make the following definition.

Definition 3. For a set X , two functions $f, g : X \rightarrow X$ are *dynamically equivalent* if there exists a bijection $\sigma : X \rightarrow X$ such that $f = \sigma^{-1}g\sigma$.

Dynamical equivalence of two maps means that they induce the same dynamics on X up to a relabeling of its elements. It is easy to check that dynamical equivalence is an equivalence relation. Let $D(n)$ denote the number of dynamical equivalence classes of the set of quadratic polynomial mappings from k to k with coefficients in k . Our problem is to estimate $D(n)$. It follows from Theorem 1 that $D(n) \leq 2^{n+1}$, but this is far from optimal. We will show the following:

Theorem 4. $D(n) = 2^{O(\frac{n}{\log \log n})}$.

2. PRELIMINARY ESTIMATES FOR $D(n)$

First we note that the maps in family (1) are never dynamically equivalent to maps in family (2), with the exception of $x \mapsto x^2 + x$, which lies in both. The map $f(x) = x^2 + \beta x$ has two fixed points (namely, 0 and $\beta + 1$) but the map $g(x) = x^2 + \beta x + (\beta^2 + 1)\zeta$ has none unless $\beta = 1$, because if $g(x) = x$ for some $x \in k$, then

$$x^2 + (\beta + 1)x + (\beta^2 + 1)\zeta = 0.$$

Using the change of variables $y = \frac{x}{\beta+1}$ again,

$$(y^2 + y + \zeta)(\beta^2 + 1) = 0.$$

But $y^2 + y + \zeta = 0$ has no solution in k as $\text{tr}(\zeta) = 1$, so g only has a fixed point when $\beta^2 + 1 = 0$, i.e. $\beta = 1$ and $g(x) = x^2 + x$.

In Theorem 1 we only allowed conjugation by the subgroup of those permutations of \mathbb{F}_{2^n} that can be represented as $\psi(x) = ax + b$ with $a, b \in k$. We now allow any permutation in the subgroup generated by these maps and by the Frobenius $\sigma : x \mapsto x^2$. This yields the following:

Corollary 5. $D(n) \leq 2G(n) - 1$, where $G(n) = \frac{1}{n} (2^n + O(2^{n/2}))$ is the number of orbits of the action of $\text{Gal}(k/\mathbb{F}_2)$ on k .

Proof. By Theorem 1, every quadratic polynomial over k is equivalent to one of the form $f(x) = x^2 + \beta x$ or $g(x) = x^2 + \beta x + (\beta^2 + 1)\zeta$. Let $\sigma(x) = x^2$ be the Frobenius, which generates $\text{Gal}(k/\mathbb{F}_2)$, and let $\psi(x) = x + \zeta(\sigma(\beta) + 1)$. We compute

$$\begin{aligned} (\sigma f \sigma^{-1})(x) &= x^2 + \sigma(\beta)x \\ ((\sigma^{-1}\psi)^{-1}g\sigma^{-1}\psi)(x) &= x^2 + \sigma(\beta)x + (\sigma(\beta)^2 + 1)\zeta. \end{aligned}$$

This shows that the dynamical equivalence class of $f(x)$ and $g(x)$ is unchanged by replacing β by $\sigma(\beta)$, or by anything in the Galois orbit

of β . The map $x \mapsto x^2 + x$ is the only one that lies in both families (1) and (2), giving the upper bound of $2G(n) - 1$. \square

3. AFFINE-LINEAR DYNAMICS ON $(\mathbb{F}_q)^n$

To prove Theorem 4 we use linear algebra. Quadratic polynomial maps from k to itself are affine-linear maps of k as an \mathbb{F}_2 -vector space. Proposition 9 decomposes such a map into a linear map and an affine map whose dynamical equivalence class is determined by an integer partition, then Proposition 11 gives a sufficient condition for dynamical equivalence of linear maps that depends on the factorization of their characteristic polynomials. Counting characteristic polynomials yields Theorem 13, which is an upper bound on the number of dynamical equivalence classes of affine-linear maps of $(\mathbb{F}_q)^n$.

The following useful definition is taken from [11] and will play a crucial role in our counting argument.

Definition 6. The *order* of $f \in \mathbb{F}_q[x]$ with $f(0) \neq 0$ is the smallest positive n such that $f \mid x^n - 1$. We write $\text{ord } f$ for the order of f .

Proposition 7. Let $O_q(k)$ be the number of integers that occur as orders of irreducible polynomials over \mathbb{F}_q of degree k . Then

$$O_q(k) = \sum_{d \mid k} \tau(q^d - 1) \mu(k/d).$$

Proof. For $f \in \mathbb{F}_q[x]$ irreducible of degree k with $f(0) \neq 0$, it is easy to show that $\text{ord } f$ is the multiplicative order of the element $[x]$ in the field $\mathbb{F}_q[x]/(f) \cong \mathbb{F}_{q^k}$ [11]. Each nonzero $\alpha \in \mathbb{F}_{q^k}$ has an irreducible minimal polynomial over \mathbb{F}_q of degree dividing k , and the order of this polynomial is the multiplicative order of α . All divisors of $|\mathbb{F}_{q^k}^\times| = q^k - 1$ occur as multiplicative orders of some $\alpha \in \mathbb{F}_{q^k}$, and all irreducible polynomials over \mathbb{F}_q of degree dividing k split in \mathbb{F}_{q^k} , so this proves

$$\sum_{d \mid k} O_q(d) = \tau(q^k - 1)$$

and the proposition follows by Möbius inversion. \square

Lemma 8. Let V and W be vector spaces, and let the pairs of linear maps $A, B : V \rightarrow V$ and $C, D : W \rightarrow W$ be dynamically equivalent, that is, $A = \phi^{-1}B\phi$ and $C = \psi^{-1}D\psi$ for some bijections $\phi : V \rightarrow V$ and $\psi : W \rightarrow W$. Then $M = A \oplus C$ and $N = B \oplus D$ are dynamically equivalent linear maps from $V \oplus W$ to itself.

Proof. Define $\theta : V \oplus W \rightarrow V \oplus W$ as ϕ on V and ψ on W , extending linearly to $V \oplus W$. Then $M = \theta^{-1}N\theta$. \square

Proposition 9. *Let V be a finite dimensional vector space over \mathbb{F}_q . Let $T : V \rightarrow V$ be defined as $Tx = Ax + b$, where $A : V \rightarrow V$ is linear and $b \in V$. If there exists $s \in V$ such that $Ts = s$, then T is dynamically equivalent to A . If no such s exists, then there is a direct sum decomposition $V = V_1 \oplus V_2$, $T = T_1 \oplus T_2$, $T_1x = A_1x + b_1$ and $T_2x = A_2x + b_2$, such that*

- (1) *There exists some $s \in V_2$ with $T_2s = s$, so T_2 is dynamically equivalent to A_2 .*
- (2) *There exists a direct sum decomposition $V_1 = W_1 \oplus \cdots \oplus W_k$, with $T_1 = S_1 \oplus \cdots \oplus S_k$, $S_i : W_i \rightarrow W_i$ affine-linear, such that the dynamical equivalence class of S is determined entirely by the $\dim W_i$, which form a partition of $\dim V_1$.*

Proof. This is proved in [18], where the notion of an identical transition graph is the same as ours of dynamical equivalence. There it is shown that, if we let $t_i = \dim W_i$, each S_i induces $\frac{q^{t_i}}{\text{ord}(x-1)^{t_i+1}}$ cycles of length $\text{ord}(x-1)^{t_i+1}$ on W_i . By Lemma 8, the set $\{\dim W_i\}$ determines the dynamical equivalence class of S . See [5] for more about transition graphs of linear maps of vector spaces over finite fields. \square

Lemma 10. *Let V be a finite dimensional \mathbb{F}_q -vector space. Let the linear map $A : V \rightarrow V$ have characteristic polynomial f^r , where f is irreducible over \mathbb{F}_q and $f(0) \neq 0$. The dynamical equivalence class of A is determined by $\deg f$, $\text{ord } f$, and an integer partition of r .*

Proof. There exist direct sum decompositions $V = V_1 \oplus \cdots \oplus V_m$ and $A = A_1 \oplus \cdots \oplus A_m$ such that in some basis of each V_i , $A_i : V_i \rightarrow V_i$ can be written as the companion matrix of f^{λ_i} for some λ_i , and $\sum \lambda_i = r$.

The orbits in V_i under the map A_i are cycles that correspond to all linearly recurrent sequences over \mathbb{F}_q with characteristic polynomial f^{λ_i} , and the lengths of the cycles are the periods of these sequences [11]. The integers that occur as periods of these sequences and the number of sequences with each period can be explicitly computed from the data $\text{ord } f$, $\deg f$, and λ_i [11, Theorem 6.63]. This determines the dynamical equivalence class of each A_i , and by Lemma 8, the dynamical equivalence class of A . \square

Proposition 11. *Let V be a finite dimensional vector space over \mathbb{F}_q and let $A : V \rightarrow V$ be a linear map. Let $p \in \mathbb{F}_q[x]$ be the characteristic*

polynomial of A and write its factorization into irreducibles as

$$p = x^{r_0} \prod_{i=1}^m p_i^{r_i}$$

where the p_i are distinct and no p_i equals x . The dynamical equivalence class of A is completely determined by an integer partition of each r_i and two lists of m positive integers: $\{\deg p_i\}$ and $\{\text{ord } p_i\}$.

Proof. By the theory of the Jordan canonical form [6] there exist direct sum decompositions $A = A_0 \oplus \cdots \oplus A_m$ and $V = V_0 \oplus \cdots \oplus V_m$ where $A_i : V_i \rightarrow V_i$, the characteristic polynomial of A_0 is x^{r_0} , and the characteristic polynomial of A_i is $p_i^{r_i}$ for $i \geq 1$. The Jordan form of the nilpotent map A_0 is specified by a partition of r_0 in which each part is the size of a Jordan block, so this partition determines the similarity class of A_0 and hence the dynamical equivalence class (if two linear maps are similar, they are dynamically equivalent).

Suppose that $\deg p_i$ and $\text{ord } p_i$ are given for $i \geq 1$. Specifying a partition of each r_i determines the dynamical equivalence class of each A_i by Lemma 10, which in turn determines the dynamical equivalence class of A by Lemma 8. \square

Before proceeding with the proof of Theorem 13, we record a proposition that will be needed at a key moment in the counting argument.

Proposition 12. *Let $i_q(n)$ denote the maximum possible number of distinct irreducible factors of a degree n polynomial over \mathbb{F}_q . Then*

$$i_q(n) = O\left(\frac{n}{\log n}\right).$$

Proof. For $q \geq 3$ it is proved in [9, Lemma A1] that

$$i_q(n) \leq \frac{n}{\log_q(n) - 3},$$

which immediately implies the proposition when $q \neq 2$. As we only require a weaker big- O estimate, we present a simplified version of the proof in [9] which also works for $q = 2$.

For $f \in \mathbb{F}_q[x]$, let $\omega(f)$ denote the number of distinct irreducible factors of f . We construct f such that $\omega(f) = i_q(n)$ by a greedy algorithm. That is, first multiply together all degree 1 irreducibles, then all degree 2 irreducibles, and so on, until multiplying f by another irreducible would raise its degree higher than n . Then $\deg f \leq n$ and no degree n polynomials have more distinct irreducible factors than f .

It suffices to prove the proposition for polynomials of the form

$$f = g \left(\prod_{\substack{p \text{ irreducible} \\ \deg p < k}} p \right)$$

where g is a product of m irreducible polynomials of degree k , each of which appears with multiplicity 1. Let $I_q(j)$ denote the number of irreducible polynomials over \mathbb{F}_q of degree j . We have

$$(3.1) \quad \omega(f) = \sum_{j=1}^{k-1} I_q(j) + m$$

and

$$(3.2) \quad \deg(f) = \sum_{j=1}^{k-1} j I_q(j) + mk \leq n.$$

We now show that the inequality

$$(3.3) \quad \sum_{j=1}^k I_q(j) \leq \frac{3}{k} \sum_{j=1}^k j I_q(j)$$

holds for large k . As $I_q(j) \leq q^j/j$ [11] we have

$$\begin{aligned} \sum_{j=1}^k I_q(j) &\leq \sum_{j=1}^k \frac{q^j}{j} \leq \sum_{1 \leq j \leq k/2} q^j + \sum_{k/2 < j \leq k} \frac{q^j}{k/2} \\ &\leq \frac{q^{k/2+1} - q}{q - 1} + \frac{2q^{k+1} - 2q^{k/2}}{k(q - 1)} \\ &\leq \frac{1}{q - 1} \left(q^{k/2+1} + \frac{2q^{k+1}}{k} \right). \end{aligned}$$

As $I_q(j) \geq q^j/j - q^{j/2+1}/(j(q - 1))$ [11] we have

$$\begin{aligned} \frac{3}{k} \sum_{j=1}^k j I_q(j) &\geq \frac{3}{k} \left(\sum_{j=1}^k q^j - \frac{q^{j/2+1}}{q - 1} \right) \\ &= \frac{1}{q - 1} \left(\frac{3q^{k+1} - 3q}{k} - \frac{3q^{(k+1)/2+1} - 3q^{3/2}}{k(q - 1)} \right). \end{aligned}$$

Therefore equation 3.3 holds if

$$q^{k/2+1} + \frac{2q^{k+1}}{k} \leq \frac{3q^{k+1} - 3q}{k} - \frac{3q^{(k+1)/2+1} - 3q^{3/2}}{k(q - 1)}.$$

or equivalently

$$q^{k/2+1} + \frac{3q}{k} + \frac{3q^{(k+1)/2+1}}{k(q-1)} \leq \frac{q^{k+1}}{k} + \frac{3q^{3/2}}{k(q-1)}.$$

Comparing powers of q on both sides, it is clear that this inequality holds for large k .

Returning to equations 3.1 and 3.2, we use 3.3 to conclude

$$i_q(n) = \omega(f) = \sum_{j=1}^{k-1} I_q(j) + m \leq \frac{3}{k} \left(\sum_{j=1}^{k-1} j I_q(j) + mk \right) \leq \frac{3n}{k}.$$

It only remains to show $k \geq C \log n$ for some C . By our construction of f , the largest that n can be for a given k occurs when g is the product of all degree k irreducible polynomials over \mathbb{F}_q . For this g , $n \leq k + \sum_{j=1}^k j I_j$ because if n exceeded this amount, we could add a degree $k+1$ irreducible factor to f . So

$$n \leq k + \sum_{j=1}^k j I_j \leq k + \sum_{j=1}^k q^k \leq k + \frac{q^{k+1} - q}{q-1} \leq k + q^{k+1}.$$

When k is large, $k < q^{k+1}$. Recall that $q \geq 2$. These imply that $n \leq k + q^{k+1} \leq q^{k+2}$, so $\log_q(n) \leq k+2$. For $n \geq q^4$

$$k \geq \log_q(n) - 2 \geq \frac{1}{2} \log_q(n),$$

which completes the proof. \square

We now combine Propositions 9 and 11 to give an upper bound on the number of dynamical equivalence classes of affine-linear maps from $(\mathbb{F}_q)^n$ to itself. The upper bound of dynamical equivalence classes of quadratic polynomial maps on \mathbb{F}_{2^n} will be an immediate corollary.

Theorem 13. *Let $V = (\mathbb{F}_q)^n$. Let $E(n)$ denote the number of dynamical equivalence classes of affine-linear maps from V to itself. Then*

$$E(n) = \exp \left(O \left(\frac{n}{\log \log n} \right) \right).$$

Remark 14. The upper bound on $E(n)$ in Theorem 13 depends on the existence of nonlinear conjugacies between affine-linear maps. Compare this bound with the well-known fact that the number of conjugacy classes in $\text{GL}(n, q)$ is $q^n + O(q^{\lfloor \frac{n-1}{2} \rfloor})$ [17].

Proof. Let $Tx = Ax + b$ be an affine-linear map from V to itself, where A is linear and $b \in V$. By Proposition 9, $V = V_1 \oplus V_2$ and $T = T_1 \oplus T_2$ such that the dynamical equivalence class of T_1 is determined by a

partition of $\dim V_1$ and T_2 is equivalent to a linear map A_2 . (It may be the case that $V_1 = 0$ and $T_1 = 0$.) Let

$$p = x^{r_0} \prod_{i=1}^m p_i^{r_i}$$

be the characteristic polynomial of A_2 . Note that $\deg p \leq n$. By Proposition 11, the dynamical equivalence class of A_2 is determined by a partition of each r_i and the two lists of m integers $\{\deg p_i\}$ and $\{\text{ord } p_i\}$. We estimate the number of ways of specifying these data. Assume for the moment that the $\deg p_i$ and the partitions of the r_i are given and that we need to assign orders to the p_i .

Let $d_i = \deg p_i$. By Proposition 7, the number of possible ways to assign the $\text{ord } p_i$ is

$$\prod_{i=1}^m O_q(d_i) \leq \prod_{i=1}^m \tau(q^{d_i} - 1).$$

We split this into two products over the ranges $d_i < d$ and $d_i \geq d$ for some d to be chosen later. First we estimate the quantity

$$C_1 = \prod_{d_i < d} \tau(q^{d_i} - 1).$$

Using the trivial estimate $\tau(x) < x + 1$,

$$C_1 < \prod_{d_i < d} q^{d_i} = q^{\sum_{d_i < d} d_i}.$$

Each d_i is the degree of a distinct irreducible polynomial over \mathbb{F}_q . As in the proof of Proposition 12, $I_q(k) \leq \frac{q^k}{k}$, so

$$\sum_{d_i < d} d_i \leq \sum_{k=1}^{d-1} k I_k \leq \sum_{k=1}^{d-1} q^k \leq q^d.$$

Therefore $C_1 < q^{q^d}$.

Now we estimate

$$C_2 = \prod_{d_i \geq d} \tau(q^{d_i} - 1).$$

By the estimate on $\tau(x)$ in [4, Theorem 8.8.9], there exists c such that

$$\tau(q^{d_i} - 1) \leq 2^{\frac{c \log(q^{d_i} - 1)}{\log \log(q^{d_i} - 1)}}.$$

This implies

$$C_2 \leq \prod_{d_i \geq d} 2^{\frac{cd_i \log q}{\log \log(q^{d_i} - 1)}} \leq 2^{\frac{c \log q \sum_{d_i \geq d} d_i}{\log \log q^{d-1}}}$$

where we use the inequality $q^d - 1 \geq q^{d-1}$, which is true for $q \geq 2$ and $d \geq 1$. Also, $\sum_{d_i \geq d} d_i \leq n$, so

$$C_2 \leq 2^{\frac{c(\log q)n}{\log(d-1) + \log \log q}}.$$

Putting these estimates together we have

$$\prod_{i=1}^m \tau(q^{d_i} - 1) = C_1 C_2 \leq \exp \left(q^d \log q + \log 2 \frac{c(\log q)n}{\log(d-1) + \log \log q} \right).$$

Choose $d = \frac{\log n}{2 \log q}$ and note $\log(d-1) \geq \log(d/2)$ for $d \geq 1$. Then

$$\begin{aligned} q^d \log q + \log 2 \frac{c(\log q)n}{\log(d-1) + \log \log q} &\leq n^{1/2} \log q + \frac{cn \log 2 \log q}{\log \frac{\log n}{4 \log q} + \log \log q} \\ &= O \left(\frac{n}{\log \log n} \right) \end{aligned}$$

Now we estimate the number of ways that the r_i and $\deg p_i$ can occur. Because $\sum_{i=1}^m r_i \deg p_i = \dim V_2$, the $\deg p_i$ form a partition of $\dim V_2$ in which each appears r_i times. This is a “factorization pattern” of $\dim V_2$ as in [7] which is specified by first picking a partition of $\dim V_2$ into parts of size k , each of which occurs s_k times, and then further dividing each s_k into parts r_i . Let $b(n)$ denote the number of factorization patterns of n . It is mentioned in [1] and proved in [13] that

$$b(n) = \exp \left(B \sqrt{n \log n} + O(\sqrt{n}) \right).$$

Finally, we need to choose a partition of each r_i and a dynamical equivalence class for T_1 given by a partition of $\dim V_1$. If $P(x)$ denotes the partition function, we have $P(x) \leq \exp(K\sqrt{x})$ [2]. The number of ways to specify all these partitions is

$$\begin{aligned} P(\dim V_1) \prod_{i=0}^m P(r_i) &\leq P(n) \exp \left(K \sum_{i=0}^m \sqrt{r_i} \right) \\ &\leq \exp \left(K \sqrt{n} + K \sqrt{m+1} \sqrt{\sum_{i=0}^m r_i} \right) \leq \exp(K \sqrt{n} (1 + \sqrt{m+1})) \\ &= \exp \left(K \sqrt{n} \left(1 + O \left(\sqrt{\frac{n}{\log n}} \right) \right) \right) = \exp \left(O \left(\frac{n}{\sqrt{\log n}} \right) \right). \end{aligned}$$

The inequality $\sum_{i=0}^m \sqrt{r_i} \leq \sqrt{m+1} \sqrt{\sum_{i=0}^m r_i}$ follows from the standard fact that the arithmetic mean of the $\sqrt{r_i}$ is at most the root mean square, and $m+1$ is the number of distinct irreducible factors of p , so $m+1 = O(n/\log n)$ by Proposition 12.

Putting this all together, the number of ways to choose a dynamical equivalence classes for T_1 and T_2 , and therefore a dynamical equivalence class for T by Lemma 8, is at most

$$b(n) \exp \left(O \left(\frac{n}{\sqrt{\log n}} \right) \right) \exp \left(O \left(\frac{n}{\log \log n} \right) \right) = \exp \left(O \left(\frac{n}{\log \log n} \right) \right)$$

which completes the proof. Theorem 4 follows immediately. \square

Remark 15. It seems possible that the estimates in Theorem 13 could be improved. The main estimate used for $\tau(q^d - 1)$ is the worst-case estimate on $\tau(x)$ that follows from the prime number theorem. It may be possible to give a better estimate based on the distribution of multiplicative orders of q modulo various integers n . (If n divides $q^d - 1$, then d is a multiple of the multiplicative order of $q \bmod n$.) Questions along these lines tend to be difficult, even for $q = 2$. See [3], [10], [12], [14], and [15] for some related work.

Acknowledgements. This research was partly supported by NSF grant no. CCF-0635355. The second author would like to thank ICERM for an invitation to the Spring 2012 Semester Program on Complex and Arithmetic Dynamics, which facilitated the development of these ideas. We would like to thank Joe Silverman for helpful comments and observations.

REFERENCES

- [1] A. K. Agarwal and G. L. Mullen. Partitions with “ $d(a)$ copies of a ”. *J. Combin. Theory Ser. A*, 48(1):120–135, 1988.
- [2] T. M. Apostol. *Introduction to analytic number theory*. Springer-Verlag, New York, 1976. Undergraduate Texts in Mathematics.
- [3] V. Arnold. Number-theoretical turbulence in Fermat-Euler arithmetics and large Young diagrams geometry statistics. *J. Math. Fluid Mech.*, 7(suppl. 1):S4–S50, 2005.
- [4] E. Bach and J. Shallit. *Algorithmic number theory. Vol. 1: Efficient algorithms*. Foundations of Computing Series. MIT Press, Cambridge, MA, 1996.
- [5] B. Elspas. The theory of autonomous linear sequential networks. In *Linear Sequential Switching Circuits*, pages 21–61. Holden-Day, San Francisco, Calif., 1965.
- [6] K. Hoffman and R. Kunze. *Linear algebra*. Second edition. Prentice-Hall Inc., Englewood Cliffs, N.J., 1971.
- [7] R. A. Hultquist, G. L. Mullen, and H. Niederreiter. Association schemes and derived PBIB designs of prime power order. *Ars Combin.*, 25:65–82, 1988.
- [8] I. M. Isaacs. *Algebra: A graduate course*. Brooks/Cole Publishing Co., Pacific Grove, CA, 1994.
- [9] Michael Kaminski and Nader H. Bshouty. Multiplicative complexity of polynomial multiplication over finite fields. *J. Assoc. Comput. Mach.*, 36(1):150–170, 1989.

- [10] P. Kurlberg and C. Pomerance. *On a problem of Arnold: the average multiplicative order of a given integer*. To appear in Algebra and Number Theory.
- [11] R. Lidl and H. Niederreiter. *Introduction to finite fields and their applications*. Cambridge University Press, Cambridge, first edition, 1994.
- [12] F. Luca and I. E. Shparlinski. Average multiplicative orders of elements modulo n . *Acta Arith.*, 109(4):387–411, 2003.
- [13] D. J. B. Mitchell. Generating functions for various sets of solid partitions. Ph. D. Thesis, Penn. State Univ., 1972.
- [14] M. R. Murty, M. Rosen, and J. H. Silverman. Variations on a theme of Romanoff. *Internat. J. Math.*, 7(3):373–391, 1996.
- [15] Carl Pomerance. On primitive divisors of Mersenne numbers. *Acta Arith.*, 46(4):355–367, 1986.
- [16] J. H. Silverman. *The arithmetic of dynamical systems*, volume 241 of *Graduate Texts in Mathematics*. Springer, New York, 2007.
- [17] Richard P. Stanley. *Enumerative combinatorics. Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2012.
- [18] K. C. Wang. Transition graphs of affine transformation on vector spaces over finite fields. *J. Franklin Inst.*, 283:55–72, 1967.

ERIC BACH, DEPARTMENT OF COMPUTER SCIENCES, UNIVERSITY OF WISCONSIN-MADISON, MADISON, WI 53706, USA

E-mail address: bach@cs.wisc.edu

ANDREW BRIDY, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN-MADISON, MADISON, WI 53706, USA

E-mail address: bridy@math.wisc.edu